University of Washington

Abstract

Certain observations concerning the effects of epistasis on complex traits and the evolution of genomes.

Maximilian Press

Co-Chairs of the Supervisory Committee:
Associate Professor Christine Queitsch
Department of Genome Sciences

Associate Professor Elhanan Borenstein
Department of Genome Sciences

The informational content of genomes is usually interpreted as a sum of one-to-one relationships between nucleotides at certain genomic positions and phenotypic outcomes. While such interpretations have the virtue of simplicity, they are often unsuccessful in elucidating the working of biological systems. Many have called for such models to explicitly consider epistasis, which can be defined as any consideration of interactions between genomic elements. In this thesis, I consider some empirical cases where epistasis may help us to understand how genomes evolve and how genotype-phenotype maps are built. In the first part of this thesis, I consider a particular case of a fast-evolving genetic element (the *ELF3* short tandem repeat in *Arabidopsis thaliana*) that shows widespread epistasis, and propose that such elements are likely to accumulate epistatic interactions by acting as mutational modifiers. This element is a polyglutamine-encoding trinucleotide in the *A. thaliana* gene *ELF3*. I go on to show some molecular mechanisms by which the element participates in epistasis, their phenotypic consequences, and make some observations on other short tandem repeats. Briefly, these observations suggest that we may be able to specifically identify such epistatic hubs among highly variable genetic elements. In the second part of

this thesis, I start with the assumption of epistasis between genes, and explore how this assumption can be used to understand the evolution of bacterial genome content. First, I take Hsp90, the known epistatic hub, and infer its coevolution with other genes through coordinated gains and losses across bacterial diversity. I further extend the underlying phylogenetic model to predict new 'clients' of bacterial Hsp90, which have remained elusive when pursued through purely experimental approaches. Collaborators were able to validate certain of these predicted clients. Last, I attempt an analogy between prokaryotic genome evolution and the much better-understood field of protein evolution. I propose that, like protein evolution by substitution, genome evolution by horizontal acquisition of genes is substantially constrained by epistasis. I go on to infer the existence of such epistatic dependencies, where one gene in an ancestral genome promotes the acquisition of a second gene. A network of such dependencies shows a chronological structuring of gene acquisitions through prokaryotic evolution, suggesting universal assembly patterns by which genomes acquire functions. I go on to show that these dependencies are taxonomically universal (i.e. not restricted to particular phyla), and that they are sufficient to make reasonably good predictions about what genes a genome will gain in the future. This predictability of genome evolution by horizontal transfer supports a major assertion of the protein evolutionists, that constraining epistasis leads to predictable evolutionary outcomes. Together, these observations indicate that the genetic architecture of traits and the content of genomes are shaped by the existence of networks of gene-gene dependencies, reflecting the complex wiring of underlying biological functions.